

PATENT COOPERATION TREATY

PCT

REC'D 06 JUN 2006


WIPO

PCT

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

(Chapter II of the Patent Cooperation Treaty)

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference RL.P53038WO	FOR FURTHER ACTION See Form PCT/IPEA/416	
International application No. PCT/EP2004/050253	International filing date (day/month/year) 04.03.2004	Priority date (day/month/year) 04.03.2004
International Patent Classification (IPC) or national classification and IPC INV. H04Q7/28		
Applicant TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)		
<p>1. This report is the international preliminary examination report, established by this International Preliminary Examining Authority under Article 35 and transmitted to the applicant according to Article 36.</p> <p>2. This REPORT consists of a total of 8 sheets, including this cover sheet.</p> <p>3. This report is also accompanied by ANNEXES, comprising:</p> <p>a. <input checked="" type="checkbox"/> sent to the applicant and to the International Bureau) a total of seven sheets, as follows:</p> <p><input checked="" type="checkbox"/> sheets of the description, claims and/or drawings which have been amended and are the basis of this report and/or sheets containing rectifications authorized by this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions).</p> <p><input type="checkbox"/> sheets which supersede earlier sheets, but which this Authority considers contain an amendment that goes beyond the disclosure in the international application as filed, as indicated in item 4 of Box No. I and the Supplemental Box.</p> <p>b. <input type="checkbox"/> (sent to the International Bureau only) a total of (indicate type and number of electronic carrier(s)) , containing a sequence listing and/or tables related thereto, in electronic form only, as indicated in the Supplemental Box Relating to Sequence Listing (see Section 802 of the Administrative Instructions).</p>		
<p>4. This report contains indications relating to the following items:</p> <p><input checked="" type="checkbox"/> Box No. I Basis of the report</p> <p><input type="checkbox"/> Box No. II Priority</p> <p><input type="checkbox"/> Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability</p> <p><input type="checkbox"/> Box No. IV Lack of unity of invention</p> <p><input checked="" type="checkbox"/> Box No. V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement</p> <p><input type="checkbox"/> Box No. VI Certain documents cited</p> <p><input type="checkbox"/> Box No. VII Certain defects in the international application</p> <p><input type="checkbox"/> Box No. VIII Certain observations on the international application</p>		
Date of submission of the demand 04.01.2006	Date of completion of this report 02.06.2006	
Name and mailing address of the international preliminary examining authority:  European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Rabe, M Telephone No. +49 89 2399-8801	



**INTERNATIONAL PRELIMINARY REPORT
ON PATENTABILITY**

International application No.
PCT/EP2004/050253

Box No. I Basis of the report

1. With regard to the **language**, this report is based on the international application in the language in which it was filed, unless otherwise indicated under this item.
- ☐ This report is based on translations from the original language into the following language , which is the language of a translation furnished for the purposes of:
- ☐ international search (under Rules 12.3 and 23.1(b))
 - ☐ publication of the international application (under Rule 12.4)
 - ☐ international preliminary examination (under Rules 55.2 and/or 55.3)
2. With regard to the **elements*** of the international application, this report is based on *(replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report)*:

Description, Pages

1, 5-7	as originally filed
2-4, 4a	received on 10.01.2006 with letter of 04.01.2006

Claims, Numbers

1-16	received on 10.01.2006 with letter of 04.01.2006
------	--

Drawings, Sheets

1/3-3/3	as originally filed
---------	---------------------

- ☐ a sequence listing and/or any related table(s) - see Supplemental Box Relating to Sequence Listing

3. ☐ The amendments have resulted in the cancellation of:
- ☐ the description, pages
 - ☐ the claims, Nos.
 - ☐ the drawings, sheets/figs
 - ☐ the sequence listing (*specify*):
 - ☐ any table(s) related to sequence listing (*specify*):
4. ☐ This report has been established as if (some of) the amendments annexed to this report and listed below had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).
- ☐ the description, pages
 - ☐ the claims, Nos.
 - ☐ the drawings, sheets/figs
 - ☐ the sequence listing (*specify*):
 - ☐ any table(s) related to sequence listing (*specify*):

* If item 4 applies, some or all of these sheets may be marked "superseded."

**INTERNATIONAL PRELIMINARY REPORT
ON PATENTABILITY**

International application No.
PCT/EP2004/050253

Box No. V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes: Claims	1-16
	No: Claims	
Inventive step (IS)	Yes: Claims	1-16
	No: Claims	
Industrial applicability (IA)	Yes: Claims	1-16
	No: Claims	

2. Citations and explanations (Rule 70.7):

see separate sheet

Reference is made to the following documents:

- D1: EP 0 584 904 A
D2: US 5 157 728 A
D3: US 2002/0173325 A1

A. Citations and explanations made in respect of paragraph V:

1. The present invention relates to a **method** of processing user speech data and to a corresponding **server node** and **mobile terminal** according to the features of respective **independent claims 1, 11 and 13**.
2. **Generally**, in the field of wireless technology, the "Push to talk Over Cellular" (PoC) service is **well known**. A PoC session is set up by a subscriber initiating the session by pressing an appropriate button on his terminal, which causes a SIP INVITE message to be sent to at least one peer terminal via a PoC server in a IP Multimedia Subsystem (IMS); upon reception by the originating terminal of a SIP 202 Accepted message from the IMS, the subscriber is able to start talking, even though the peer terminal has not yet accepted the session; this initial talk burst is buffered by the PoC server; when a SIP 200 OK message indicating acceptance of the session is received from the peer terminal, the buffered talk burst is immediately sent to the peer terminal by the PoC server.

Document **D1** discloses a trunked TDMA radio communications system, wherein time delays occurring after one transceiver ends a transmission and before other transceiver transmissions may commence are minimised; after depression of a PTT button, the transceiver receives over a control channel a working frequency and a time slot assignment; immediately thereafter, an alert tone is generated to signal to the user of the transceiver that he may begin speaking immediately, ie. in the second half of the first frame, while at the same time setup tasks commence; speech input captured in the second frame is processed and buffered in the third frame and transmitted in the first time slot of the fourth frame; speech captured in the first frame after alert tone generation up to the second frame is not used to reduce delays

between generation of the alert tone and receipt of a channel access signal.

Furthermore, document **D2** discloses a PTT system wherein delays caused by buffering of a signal containing speech are substantially eliminated; a receiver receives an input speech signal and produces an output speech signal representing the input signal after a delay intentionally introduced in the system (eg. to obtain a communication channel); a buffer controller analyses at least a portion of the input signal stored in the buffer in order to determine which parts of the signal may be removed without substantial deterioration of the signal quality; in particular, the portions of the stored speech to be removed represent silence gaps in the speech, and the duration of these gaps may be shortened by some predetermined percentage.

3. A main **disadvantage** related to the known systems is that either speech may be lost (note that in document **D1**, a fixed period (ie. one frame) of the buffered input is lost, independent on whether said period comprises speech or silence) or the reduction of delays may not be optimum (note that in document **D2**, the silence periods within a buffered signal, ie. pauses in a stored speech signal, may be shortened only by a predetermined percentage).
4. The **present invention** overcomes these disadvantages by providing a **method** of processing user speech data and a corresponding **server node** and **mobile terminal** according to the features of respective **independent claims 1, 11 and 13**.

According to the **essential features of the invention**, at a processing entity for transmission of user speech data to a participant or participants in a push to talk session, **speech data** is **analysed** to identify an **initial** period of silence **following** initiation of the push to talk session, but **prior to** receipt of a session acceptance from the or each participant; an **initial** period of silence is **removed** from the speech data **prior to** sending the speech data to a receiving terminal of the or each participant.

5. The present invention provides the **advantage** of significantly reducing the delay between session initiation and acceptance by removing identified periods of silence from speech data before a session acceptance is received; no speech is lost, and

delays of variable duration may be reduced.

6. The subject-matter of the present invention as claimed in respective independent claims 1, 11 and 13 is neither disclosed in, nor rendered obvious by the remaining **prior art documents** cited in the international search report since said documents, which merely relate to a very general state of the art of PTT systems and session management in wireless communication networks, do **not** describe or render (in combination) obvious the method of processing user speech data and the corresponding server node and mobile terminal according to the features of respective independent claims 1, 11 and 13.
7. The subject-matter of independent claims 1, 11 and 13 therefore is considered to be **new** and to **involve an inventive step**, Article 33 (2) and (3) PCT.
8. As **claims 2 to 10, 12 and 14 to 16** are dependent on respective independent claims 1, 11 and 13, said claims 2 to 10, 12 and 14 to 16 do **also meet** the requirements of Article 33 (2) and (3) PCT.
9. The present invention is **susceptible of industrial application**, Article 33 (4) PCT.

B. Further remarks made in respect of the present application:

1. To meet the requirements of Rule 6.3 (b) PCT, any independent claim should have been **correctly** cast in the **two-part form**, with those features which in combination are part of the nearest prior art being placed in the preamble.
2. **Reference signs** in parentheses should have been inserted in all the claims to increase their intelligibility, Rule 6.2 (b) PCT. This applies both to the preamble and to the characterizing portion.
3. The wording of **claim 16** should have been: "A terminal according to claim 13 or **14** ...", Article 6 PCT.

**INTERNATIONAL PRELIMINARY
REPORT ON PATENTABILITY
(SEPARATE SHEET)**

International application No.

PCT/EP2004/050253

4. To meet the requirements of Rule 5.1 (a) (ii) PCT, the documents **D1** to **D3**, which represent a relevant state of the art with regard to the present invention, should have been identified in the opening part of the description and the relevant

b
a
c
k
g
r
o
u
n
d
a
r
t
d
i
s
c
l
o
s
e
d
t
h
e
r
e
i
n
s
h

**INTERNATIONAL PRELIMINARY
REPORT ON PATENTABILITY
(SEPARATE SHEET)**

International application No.

PCT/EP2004/050253

o
u
l
d
h
a
v
e
b
e
e
n
b
r
i
e
f
l
y
d
i
s
c
u
s
s
e
d
.

EPO -DG 1

RL.P53038WO

10.01.2006

2

103

- Figure 2 illustrates certain signalling associated with setting up a PoC session across the network of Figure 1 (additional messages may also be transferred between the various nodes, although these are not shown in the Figure). A subscriber initiates a session by pressing the appropriate button on his/her terminal UE#1. This causes a SIP INVITE message to be sent to the peer terminal UE#2 via the PoC server in the IMS core, followed by the transfer of further signalling between the terminals and the IMS. As already mentioned, a key component of PoC is the near instantaneous connection of parties. Significant delays in transmitting speech are therefore to be avoided.
- 10 The time between the SIP INVITE message being sent and the IMS receiving an acceptance from the called party can be as much as 3 seconds due to fundamental properties of the network (e.g. paging, Temporary Block Flow (TBF) establishment, etc). In order to speed up the initial connection process, the initiating subscriber is therefore able to start talking upon receipt by his terminal of the SIP 202 Accepted message from the IMS (usually signalled to the initiating subscriber by the playing of a tone or "beep" on his terminal), even though the called party has not yet accepted the session. The initial talk burst may be buffered by a PoC server within the network until such time as it receives the SIP 200 OK message from the peer terminal. When that message is received, the talk burst is immediately sent to the peer terminal.
- 20 Nonetheless, the delay perceived by the called party remains significant and it is desirable to reduce the delay still further.

Summary of the Invention

- 25 The inventor of the present invention has recognised that the initiating subscriber is unlikely to begin talking for a short while after the tone has been played due both to the reaction time of the subscriber and to his/her "thinking time". In the example of Figure 2, this delay is of the order of 0.8 seconds.
- 30 According to a first aspect of the present invention there is provided a method of processing user speech data at a processing entity for transmission to a participant or participants in a push to talk session over a communications network, the method comprising:

RL.P53038WO

3

following initiation of a push to talk session, but prior to receipt by the entity of a session acceptance from the or each participant, analysing the speech data to identify an initial period of silence; and removing an initial period of silence from the speech data prior to sending the speech data to a receiving terminal of the or each other participant.

The invention is particularly applicable to removing an initial period of silence from the initial speech burst provided by the initiating party of the push to talk session. This has the effect of reducing the delay between the generation of the speech burst by the initiating subscriber and the playing of the speech burst to the or each other participant.

Preferably, said communication network is a cellular telephone network and the push to talk service is a Push to talk Over Cellular (PoC) service.

The step of analysing the speech data to identify an initial period of silence may be carried out at the terminal of the initiating party or at a node within the communication network. Similarly, the step of removing the detected period of silence from the transmitted speech data may be carried out at the terminal of the initiating party or at a node within the communication network. The network node is preferably within the IP Multimedia Subsystem (IMS) in the case where the communication network is a cellular telephone network and the push to talk service is a PoC service.

In the case where the steps of detecting and removing are done at the initiating party's terminal, the step of detecting may comprise analysing the speech data during or following recording of the data at the terminal.

Certain embodiments of the invention may comprise monitoring the audio level and commencing recording of the speech only when that level exceeds some predefined threshold. This step may be carried out at the terminal of the imitating party or at a server node within the communication network. In other embodiments of the invention, an initial period expected to contain silence is predefined, and the start of the speech data is clipped to remove the predefined period. The predefined period may be fixed, or may be adaptive based upon talk/usage patterns of the user.

RL.P53038WO

4

The step of removing an initial period of silence from the speech data may be carried out in real-time, as the speech data is received, or may be carried out by post-processing stored or buffered speech data.

5

According to a second aspect of the present invention there is provided a server node for use in a communication network offering a push to talk service to subscribers, the node comprising:

10 a receiver for receiving a speech burst from a participant in a push to talk session; and

a processor for, following initiation of a push to talk session but prior to receipt by the network of a session acceptance from a receiving participant, detecting an initial period of silence in the speech data burst and removing the detected period of silence from the speech data prior to transmission to the or each other participant in the session.

15

Preferably, said server node is arranged to be located within an IP Multimedia Subsystem of a cellular telephone communications network, the node having an interface to one or more Session Initiation Protocol (SIP) servers including a Serving Call Session Control Function (S-CSCF) server.

20

According to a third aspect of the present invention there is provided a mobile terminal for use in a communication network offering a push to talk service to subscribers, the terminal comprising:

25 a receiver for receiving speech data from a terminal user; and
a processor for, following initiation of a push to talk session but prior to receipt by the mobile terminal of a session acceptance from a receiving terminal, removing a period of silence from the speech data prior to transmission to the or each other terminal participating in the session.

30 Preferably, said mobile terminal is a wireless terminal and the communication network is a cellular telephone network offering a Push to talk Over Cellular service.

RL.P53038WO

4a

The mobile terminal may be a terminal used by said terminal user, or may be another terminal participating in the session.

Brief Description of the Drawings

5

RLP53038WO

EPO -DG 1

10. 01. 2006

8

103

Claims

1. A method of processing user speech data at a processing entity for transmission to a participant or participants in a push to talk session over a communications network,
5 the method comprising:
following initiation of a push to talk session, but prior to receipt by the entity of a session acceptance from the or each participant, analysing the speech data to identify an initial period of silence; and removing an initial period of silence from the speech data prior to sending the speech data to a receiving terminal of the or each other
10 participant.
2. A method according to claim 1, wherein said speech data is an initial speech burst provided by the initiating party of the push to talk session.
- 15 3. A method according to claim 1 or 2, wherein said communication network is a cellular telephone network and the push to talk service is a Push to talk Over Cellular service.
4. A method according to any one of the preceding claims, wherein said step of
20 analysing the speech data to identify an initial period of silence is carried out at a terminal of the initiating party or a node within the communication network.
5. A method according to any one of the preceding claims, wherein the step of removing an initial period of silence from the transmitted speech data is carried out at a
25 terminal of the initiating party or a node within the communication network.
6. A method according to claim 5, wherein the network node is a Media Resource Function node.
- 30 7. A method according to claim 5, wherein the network node is located within an IP Multimedia Subsystem (IMS).

RL.P53038WO

9

8. A method according to any one of the preceding claims and comprising monitoring the audio level to determine when speech has started.

9. A method according to any one of claims 1 to 7 and comprising predefining an initial period expected to contain silence, and clipping the start of the speech data
5 remove the predefined period.

10. A method according to claim 9, wherein the predefined period is fixed or is adapted in dependence upon subscriber behaviour.

10

11. A server node for use in a communication network offering a push to talk service to subscribers, the node comprising:

a receiver for receiving a speech burst from a participant in a push to talk session; and

15 a processor for, following initiation of a push to talk session but prior to receipt by the network of a session acceptance from a receiving participant, detecting an initial period of silence in the speech data burst and removing the detected period of silence from the speech data prior to transmission to the or each other participant in the session.

20 12. A server node according to claim 11 and being arranged to be located within an IP Multimedia Subsystem of a cellular telephone communications network, the node having an interface to one or more Session Initiation Protocol (SIP) servers including a Serving Call Session Control Function (S-CSCF) server.

25 13. A mobile terminal for use in a communication network offering a push to talk service to subscribers, the terminal comprising:

a receiver for receiving speech data from a terminal user; and

a processor for, following initiation of a push to talk session but prior to receipt by the mobile terminal of a session acceptance from a receiving terminal, removing a
30 period of silence from the speech data prior to transmission to the or each other terminal participating in the session.

RL.P53038WO

10

14. A terminal according to claim 13, the terminal being a wireless terminal and the communication network being a cellular telephone network offering a Push to talk Over Cellular service.

5 15. A terminal according to claim 13 or 14, wherein the receiver comprises means for converting speech into an analogue or digital electrical signal.

16. A terminal according to claim 13 or 1614 wherein the receiver comprises means for receiving speech data over an interface link to said communication network, the
10 speech data having been generated at a peer mobile terminal.